On the Interlacing Property for Singular Values and Eigenvalues

João Filipe Queiró
Universidade de Coimbra
Departamento de Matemática
Apartado 3008
3000 Coimbra, Portugal

Submitted by David H. Carlson

ABSTRACT

A new proof of the complete interlacing theorem for singular values is presented. The technique used works equally well to obtain the interlacing theorem for eigenvalues of hermitian matrices.

1. INTRODUCTION

Given an arbitrary, not necessarily square, complex matrix M, its *singular values* are defined as the eigenvalues of the positive semidefinite matrix $(M^*M)^{1/2}$ [or $(MM^*)^{1/2}$, which, apart from zeros, has the same eigenvalues]. Several years ago, in [9], R. C. Thompson described the exact relations between the singular values of a matrix and those of a submatrix. The relations take the form of (interlacing) inequalities connecting the two sets of numbers. (The theorem is stated below.)

Thompson gave two proofs for the necessity part, and one for the sufficiency. All his arguments depend on the celebrated interlacing inequalities relating the eigenvalues of a hermitian matrix with those of a principal submatrix. These inequalities are also known to be necessary and sufficient [2, 5, 6].

Other proofs of the necessity part of the interlacing theorem for singular values are known, e.g. using the characterization of singular values: (i) as approximation numbers (see the remark at the end of [8]); (ii) via a minimax theorem (see, e.g., [1]). The corresponding result for eigenvalues of hermitian matrices can also be obtained using similar ideas.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

We give here another proof for both the necessity and the sufficiency of the interlacing theorem for singular values. Our argument is simple and geometrical in character. Although it also works for eigenvalues of hermitian matrices, we choose to present it in the singular value setting because it is slightly simpler here.

In the last section we mention some singular value estimates and their use in the determination of the asymptotic behavior of singular values of certain matrices.

2. PROOF OF THE INTERLACING THEOREM FOR SINGULAR VALUES

In its most general form, the theorem to be proved reads as follows:

THEOREM. *Let m, n, p, q be natural numbers, with $m \geqslant p$, $n \geqslant q$. Given nonnegative real numbers $\beta_1 \geqslant \beta_2 \geqslant \cdots \geqslant \beta_{\min\{m,n\}}$, $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_{\min\{p,q\}}$, there exists an $m \times n$ complex matrix with the β_i as singular values and having a $p \times q$ submatrix with the α_i as singular values if and only if*

$$\beta_i \geqslant \alpha_i, \qquad i = 1,2,\ldots,\min\{p,q\},$$

$$\alpha_i \geqslant \beta_{i+m-p+n-q}, \qquad i = 1,2,\ldots,\min\{p+q-m, p+q-n\}.$$

We begin with some standard reductions, omitting the (easy) justifications.

(1) It is enough to prove the theorem for $m = n+1$, $p = q = n$. In this case the inequalities become $\beta_1 \geqslant \alpha_1 \geqslant \beta_2 \geqslant \alpha_2 \geqslant \cdots \geqslant \beta_n \geqslant \alpha_n$.

(2) We may restrict our attention to the case where the submatrix lies in the first n rows.

(3) We may restrict our attention to matrices of the type

$$\begin{bmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \\ z_1 & \cdots & z_n \end{bmatrix}$$

where $z_1, \ldots, z_n$ are complex numbers. We denote such a matrix by $B(z_1, \ldots, z_n)$.

(4) We may restrict our attention to the case where $z_1, \ldots, z_n$ in the matrix $B(z_1, \ldots, z_n)$ are nonnegative real numbers.

Of course, for the sufficiency, it is clear that the restrictions described in (2), (3), (4) make no difference.

(5) It is enough to prove the theorem for the case where the α_i are all distinct (i.e., $\alpha_1 > \alpha_2 > \cdots > \alpha_n$).

We shall use below the following fact:

(6) If the matrix $M = [m_{ij}]$ has singular values s_i, then

$$\sum_{i,j} |m_{ij}|^2 = \sum_i s_i^2 .$$

Proof. Both members are equal to $\operatorname{tr} M^*M$. ■

Denote by $\mathbb{R}_0^{+n}$ the nonnegative orthant in $\mathbb{R}^n$. We are concerned with a mapping $s : \mathbb{R}_0^{+n} \to \mathbb{R}_0^{+n}$ defined, for $x = (x_1, \ldots, x_n) \in \mathbb{R}_0^{+n}$, by $s(x) = (s_1(x), \ldots, s_n(x))$, where $s_1(x), \ldots, s_n(x)$ are the singular values of the matrix $B(x) = B(x_1, \ldots, x_n)$ ordered so that $s_1(x) \geqslant s_2(x) \geqslant \cdots \geqslant s_n(x)$.
Three properties of s are relevant to us:

(7) Putting $\alpha = (\alpha_1, \ldots, \alpha_n)$, we have $\|s(x)\| = (\|x\|^2 + \|\alpha\|^2)^{1/2}$.

(Here and below, $\|\cdot\|$ is the euclidean norm).
This follows from (6).

(8) s is continuous.

s is a Lipschitz function, with Lipschitz constant (with respect to the euclidean norm) at most 1. This follows from the combination of a result in [3] with an often quoted observation by Wielandt, which allows the translation of many results concerning eigenvalues of hermitian matrices into results about singular values (see, e.g., the end of [9]).

(9) Assuming $\alpha_1 > \alpha_2 > \cdots > \alpha_n$, s is one-to-one.

Proof. This can be proved as follows: Suppose $s(x) = s(y)$, with $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}_0^{+n}$. Then, by (6), we have

$$\alpha_1^2 + \cdots + \alpha_n^2 + x_1^2 + \cdots + x_n^2 = \alpha_1^2 + \cdots + \alpha_n^2 + y_1^2 + \cdots + y_n^2,$$

whence $(x_1^2 - y_1^2) + \cdots + (x_n^2 - y_n^2) = 0$. Applying the same reasoning to the successive compounds of $B(x)$ and $B(y)$ and canceling out common summands (those where only α's occur), we obtain a homogeneous system of n linear equations in $x_1^2 - y_1^2, \dots, x_n^2 - y_n^2$. The matrix of this system is seen to be the Jacobian of the n elementary symmetric functions of the α_i^2. Its determinant is easily computed, and is equal to $\prod_{i<j}(\alpha_i^2 - \alpha_j^2)$ (see [7, pp. 150–151]). Since $\alpha_1 > \cdots > \alpha_n \geqslant 0$, it follows that $x_i^2 = y_i^2$, $i = 1, \dots, n$, whence, since the numbers are nonnegative, we have $x_i = y_i$, $i = 1, \dots, n$, that is, $x = y$. ■

The proof of the theorem itself begins here. Put

$$\Delta = [\alpha_1, +\infty) \times [\alpha_2, \alpha_1] \times \cdots \times [\alpha_n, \alpha_{n-1}].$$

Then the necessity part reads $s(\mathbb{R}_0^{+n}) \subseteq \Delta$, and the sufficiency reads $s(\mathbb{R}_0^{+n}) \supseteq \Delta$. Therefore, to prove the theorem we have to show that $s(\mathbb{R}_0^{+n}) = \Delta$.

We proceed by induction on n.

For $n = 1$ the situation is clear: the range of $(x_1^2 + \alpha_1^2)^{1/2}$ is $[\alpha_1, +\infty)$.

For general $n \geqslant 2$, the induction hypothesis implies that $s(\partial \mathbb{R}_0^{+n}) = \partial\Delta$, where the symbol ∂ denotes boundary. Since s is continuous and one to one, it follows that either $s(\mathbb{R}_0^{+n}) \subseteq \Delta$ or $s(\mathbb{R}_0^{+n}) \subseteq \mathbb{R}_0^{+n} \setminus \operatorname{int}\Delta$ ("int" stands for interior).

For each $r \geqslant 0$, let $\sigma_r = \{v \in \mathbb{R}^n : \|v\| = r\}$, $\sigma_r^+ = \sigma_r \cap \mathbb{R}_0^{+n}$. Also put $r' = (r^2 + \|\alpha\|^2)^{1/2}$.

To prove that $s(\mathbb{R}_0^{+n}) = \Delta$ it will be enough to show that, for every r, $s(\sigma_r^+) = \sigma_{r'}^+ \cap \Delta$, since $\mathbb{R}_0^{+n} = \bigcup_r \sigma_r^+$ and $\Delta \subseteq \bigcup_r \sigma_{r'}^+$.

By property (7) we have that $s(\sigma_r^+) \subseteq \sigma_{r'}^+$, and by the alternative above we know that either $s(\sigma_r^+) \subseteq \sigma_{r'}^+ \cap \Delta$ or $s(\sigma_r^+) \subseteq \sigma_{r'}^+ \setminus \operatorname{int}\Delta$.

Since σ_r^+ is compact, σ_r^+ and $s(\sigma_r^+)$ are homeomorphic, whence, considered as subsets of σ_r and $\sigma_{r'}$, respectively, they satisfy $s(\partial\sigma_r^+) = \partial s(\sigma_r^+)$. On the other hand, it is easily seen that $s(\partial\sigma_r^+) = \sigma_{r'}^+ \cap \partial\Delta$.

From the identity $\partial s(\sigma_r^+) = \sigma_{r'}^+ \cap \partial\Delta$ we infer that $s(\sigma_r^+) \subseteq \sigma_{r'}^+ \cap \Delta$ [otherwise $s(\sigma_r^+)$ would have boundary points elsewhere]. This already proves the necessity part of the theorem.

Now $A = s(\sigma_r^+)$ and $B = \sigma_{r'}^+ \cap \Delta$, considered as subsets of $\sigma_{r'}$, satisfy $A \subseteq B$, $\partial A = \partial B$, $\operatorname{int} A$ nonempty, and $\operatorname{int} B$ connected. This ensures that $A = B$, and the proof is complete.

3. ESTIMATES FOR SINGULAR VALUES

Let $s_1 \geqslant s_2 \geqslant \cdots$ be the singular values of a matrix $M = [m_{ij}]$. Then s_1 is equal to the spectral norm of M:

$$s_1 = \|M\| = \sup_{\|v\|=1} \|Mv\|.$$

From our knowledge of the singular values of $C_k(M)$, the successive compounds of M, we deduce that

$$s_k = \frac{\|C_k(M)\|}{\|C_{k-1}(M)\|}, \qquad k = 1,\ldots,\operatorname{rank} M$$

[we put $\|C_0(M)\| = 1$].

Suppose now that we take some estimates for the norm of a matrix:

$$l(M) \leqslant \|M\| \leqslant u(M) \tag{10}$$

[$l(M)$ and $u(M)$ may be other norms, or generalized norms; preferably they should be quantities easily computed in terms of the entries of the matrix]. We then have the following estimates for the singular values:

$$\frac{l(C_k(M))}{u(C_{k-1}(M))} \leqslant s_k \leqslant \frac{u(C_k(M))}{l(C_{k-1}(M))}, \qquad k = 1,\ldots,\operatorname{rank} M. \tag{11}$$

How good these estimates are obviously depends on the sharpness of (10). Anyway, they can hardly be of numerical interest, because of the computations needed just to write down the compounds. But they might be used for theoretical purposes. As an illustration, we determine the asymptotic behavior of the singular values of the matrices $B(x_1,\ldots,x_n)$ that appear in the previous section: For $i = 1,\ldots,n$ we have

$$\lim_{x_i \to +\infty} s(x_1,\ldots,x_n) = (+\infty, \alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_n).$$

(While $x_i \to +\infty$ the other coordinates are to remain fixed.) The proof of this, which does not need the interlacing inequalities, may be obtained by using (11) with, for example, the following rather crude bounds for the spectral norm:

$$l(M) = \max_{i,j} |m_{ij}|,$$

$$u(M) = \left[\left(\max_i \sum_j |m_{ij}|\right)\left(\max_j \sum_i |m_{ij}|\right)\right]^{1/2}.$$

(That $u(M) \geqslant \|M\|$ is shown in [4, p. 212–213].)

REFERENCES

1 D. Carlson, Minimax and interlacing theorems for matrices, *Linear Algebra Appl.* 54:153–172 (1983).
2 Ky Fan and G. Pall, Imbedding conditions for hermitian and normal matrices, *Canad. J. Math.* 9:298–304 (1957).
3 A. J. Hoffman and H. Wielandt, The variation of the spectrum of a normal matrix, *Duke Math. J.* 20:37–39 (1953).
4 P. Lancaster, *Theory of Matrices*, Academic, New York, 1969.
5 L. Mirsky, Matrices with prescribed characteristic roots and diagonal elements, *J. London Math. Soc.* 33:14–21 (1958).
6 G. N. de Oliveira, Matrices with prescribed characteristic polynomial and a prescribed submatrix, *Pacific J. Math.* 29:653–661 (1969).
7 O. Perron, *Algebra. I*, Walter de Gruyter and Co., Berlin, 1927.
8 J. F. Queiró, Invariant factors as approximation numbers, *Linear Algebra Appl.* 49:131–136 (1983).
9 R. C. Thompson, Principal submatrices IX: Interlacing inequalities for singular values of submatrices, *Linear Algebra Appl.* 5:1–12 (1972).

Received 6 March 1984